مدل‌سازی و تحلیل بازی معمای زندانی تکراری به کمک شبکه عصبی مصنوعی پادانتشار گراسبرگ

غلامعلی منتظر، نجمه رستگار رامشه و علیرضا عسکرزاده

چکیده: در اکثر اوقات، تصمیم‌گیری مؤثر در موقعیت‌های راهبردی همچون مسایل رقابتی به نگاشت غیر خطی بین محرک و پاسخ نیاز دارد. شبکه‌های عصبی مصنوعی می‌توانند در مدل‌سازی و حل این مسایل رهیافت مناسبی باشند. بازی معمای زندانی از معروف‌ترین بازی‌های مطرح‌شده در نظریه بازی‌ها است که به کمک آن بسیاری از مسایل رقابتی تحلیل می‌شود و تصمیم‌گیری را تسهیل می‌کند. در این مقاله سعی بر آن است که بازی معمای زندانی تکراری به کمک شبکه عصبی مصنوعی مدل‌سازی و تحلیل شود و به همین دلیل شبکه عصبی پادانتشار گراسبرگ برای انجام این بازی طراحی شده است. نتایج، نشان‌دهنده توانمندی این روش در مدل‌سازی کامل بازی است. نتایج حاصل از به‌کارگیری این ایده با دو روش دیگر (راهبرد TFT و مدل‌سازی با شبکه پرسپترون) نشان از کارایی محرز روش جدید است.

***کلید واژه:* بازی معمای زندانی، راهبرد TFT، شبکه پادانتشار گراسبرگ، شبکه عصبی مصنوعی.**

۱- مقدمه

در دنیای رقابتی تجارت، یکی از مهم‌ترین مسایل، مطالعه و حدس فعالیت‌ها و کنش‌های رقابتی رقیب است. اگر اطلاعات درباره راهبردهای رقبا قابل دسترس باشد، حداکثرکردن درآمد شرکت با در نظر گرفتن و به حساب آوردن حرکات رقیب امکان‌پذیر می‌شود. نظریه بازی[۱] (GT) برای حل چنین مسایلی به‌وجود آمده است، با این فرض که در یک موقعیت رقابتی، شرکت‌کنندگان به طرز معقولی عمل می‌کنند و سعی می‌کنند تضادهای[۲] رقابتی را به نفع خودشان حل و فصل کنند [۱].

نظریه بازی در اواخر قرن نوزدهم مطرح شد ولی اولین بار این نظریه را جان فون نیومن[۳] در سال ۱۹۴۴ میلادی طی جنگ جهانی دوم برای تحلیل تضادهای سیاسی و رقابت‌های اقتصادی ارائه کرد. دیدگاه بنیادی نظریه بازی‌ها به‌کارگیری منطق بازی برای مدل‌کردن زندگی واقعی است که در آن بازیکنان در مقابل راهبردهای یکدیگر به رقابت می‌پردازند [۲]. نظریه بازی‌ها در زمینه‌های متفاوتی از جمله اقتصاد [۳]، امنیت شبکه [۴]، قیمت‌گذاری در اینترنت [۵]، مسیریابی در شبکه‌های اقتضایی[۴] [۶]، مدیریت زنجیره تأمین [۷]، زیست‌شناسی [۸] و بسیاری از زمینه‌های دیگر کاربرد دارد.

نظریه بازی‌ها، مطالعه تعاملات بین تصمیم‌گیرندگان مستقل از یکدیگر است و تعاملات بین تصمیم‌گیرندگان، مفهوم بازی در نظریه بازی‌ها را ایجاد می‌کند [۹]. این که تصمیم‌گیرندگان در حین بازی عاقلانه رفتار می‌کنند و به دنبال بیشینه‌کردن سود خود هستند و در اتخاذ این تصمیم رفتار دیگر تصمیم‌گیرندگان را در نظر می‌گیرند، فرض اساسی نظریه بازی‌ها است. در نظریه بازی‌ها رکن اصلی بازی، بازیکن[۵] است. برای هر بازیکن تعداد محدودی راهبرد[۶] وجود دارد که به آنها "مجموعه راهبرد بازیکن" گفته می‌شود و از هر راهبرد نتیجه‌ای حاصل می‌شود که همان عایدی (پیامد)[۷] بازیکن است. عایدی هر بازیکن یکی از عناصر مهم بازی است. راه حل، شرح روش‌مند[۸] نتایجی است که ممکن است در یک بازی رخ دهد. یکی از معمول‌ترین مفاهیم بازی در نظریه بازی‌ها تعادل نش[۹] است. در تعادل نش تمامی بازیکنان بهترین راهبرد خود را با توجه به رفتار بازیکنان مقابل انتخاب می‌کنند [۱۰]. یافتن نقطه (یا نقاط) تعادل نش در بازی همیشه مشکلاتی را در برداشته است. در برخی بازی‌ها وقتی بازیکنان راهبرد خالص (انتخاب راهبرد با احتمال صفر یا یک) را بر می‌گزینند، تعادل نش وجود ندارد و لازم است محاسبات زیادی برای پیداکردن تعادل نش ترکیبی (راهبردها با احتمال بین صفر و یک انتخاب می‌شوند) انجام شود. در برخی دیگر از بازی‌ها، بیش از یک تعادل نش وجود دارد و در این حالت، بازیکن برای تصمیم‌گیری بین چندین تعادل نش دچار مشکل می‌شود. برای حل این مشکلات روش‌های هوشمندی برای تحلیل بازی و پیداکردن تعادل نش پیشنهاد شده است که شامل روش‌های زیر است:

۱-۱ نظریه بازی‌های تکاملی

در نظریه بازی‌های تکاملی[۱۰]، بازیکنان بازی و راهبردهایشان باید با جوامعی از بازیکنان شبیه‌سازی شوند و بدین منظور فرایندی برای تکامل جامعه تعریف می‌شود. تحت این فرایند، راهبردهای جوامع بازیکنان به مقداری پایدار همگرا می‌شود که این مقدار همان تعادل نش بازی است [۱۱].

به غیر از تکرار[۱۱] که نوعی پویایی را نشان می‌دهد، تقلید[۱] نوع دیگری

این مقاله در تاریخ ۱۵ اسفند ماه ۱۳۸۹ دریافت و در تاریخ ۱۸ مهر ماه ۱۳۹۱ بازنگری شد.

غلامعلی منتظر، گروه مهندسی فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران، (email: montazer@modares.ac.ir).

نجمه رستگار رامشه، گروه مهندسی فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران، (email: samane.rastegar@gmail.com).

علیرضا عسکرزاده، پژوهشکده انرژی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فنّاوری پیشرفته، کرمان، (email: a.askarzadeh@kgut.ac.ir).

1. Game Theory
2. Conflicts
3. John Von Neumann
4. Ad hoc Network
5. Player
6. Strategy
7. Payoff
8. Systematic
9. Nash Equilibrium
10. Evolutionary Game
11. Replication

Archive of SID

www.SID.ir

از پویایی است که به‌طور وسیعی در این نظریه مورد مطالعه قرار گرفته است. در تقلید، هر بازیکن در ابتدا راهبرد خالصی را بازی می‌کند اما با گذر زمان، بازیکنان با یکدیگر تعامل برقرار می‌کنند. بعد از این تعامل دونفره، اگر بازیکنی ببیند بازیکن دیگر عایدی بیشتری به‌دست آورده، تلاش می‌کند با اتخاذ راهبرد او، رفتار بهتر را تقلید کند. معمولاً مدل‌سازی و شبیه‌سازی بازی‌های تکاملی به کمک الگوریتم ژنتیک صورت می‌گیرد [۱۲].

۱-۲ نظریه بازی‌های فازی[۲]

در این رویکرد اولویت میان راهبردهای منتخب به‌صورت فازی در نظر گرفته می‌شود و بدین ترتیب صورت‌بندی جدیدی از مسئله در قالب منطق فازی شکل می‌گیرد. چاکری[۳] نوعی بازی را مطرح کرده است که برای حل آن با استفاده از رابطه اولویت قطعی، بیش از یک تعادل نش پدید می‌آید ولی با استفاده از رابطه اولویت فازی، تنها یک تعادل نش (که بهترین تعادل نش بازی است) به‌دست می‌آید. بنابراین تصمیم‌گیرنده (بازیکن) به راحتی راهبرد خود را انتخاب می‌کند [۱۳]. در رویکردی دیگر گارازیک[۴] و کروز[۵] سیستمی فازی را پیشنهاد داده‌اند که در آن ابتدا راهبردهای بازی فازی می‌شوند، سپس ماتریس اولویت فازی با استفاده از قواعد "اگر- آنگاه" در مرحله استنتاج[۶] تولید می‌شود و سرانجام پس از نافازی‌سازی[۷]، بازی قطعی و تعادل نش به‌دست می‌آید [۱۴].

۱-۳ شبکه عصبی مصنوعی

استفاده از شبکه عصبی مصنوعی رویکرد دیگری است که برای تحلیل بازی به‌کار گرفته شده است [۱۵]. دلیل استفاده از شبکه‌های عصبی مصنوعی برای تحلیل بازی این است که بازیکنان ممکن است به دلایل مختلفی همچون کمبود وقت یا تعدد راهبردها نتوانند در تمامی شرایط بازی، امکان حضور در صحنه و انجام بازی را داشته باشند. در این صورت استفاده از شبکه عصبی راهکار مناسبی است که می‌تواند به‌طور مؤثری با یادگیری قواعد بازی و مدل‌سازی شرایط آن، اجرای بازی را بر عهده گیرد. از سوی دیگر شبکه با یادگیری قواعد و محیط بازی می‌تواند در پیش‌بینی و حدس درست راهبرد حریف، کمک بسیاری به تصمیم‌گیرنده در اتخاذ تصمیم بهینه و بیشینه‌کردن عایدی داشته باشد. البته باید اشاره کرد که کاربرد این رویکرد در مدل‌سازی بازی جدیدتر از سایر روش‌هاست و به همین دلیل تعداد پژوهش‌های مرتبط، انگشت‌شمار است. در این میان می‌توان به نتایج تحقیقات لشنو[۸] و همکارانش اشاره کرد که شرایط اقتصادی بازار را به شکل بازی مدل کرده‌اند و از شبکه پرسپترون برای تحلیل آن بهره جسته‌اند [۱۶]. در تحقیقی دیگر کوراد[۹] و لیو[۱۰] شبکه عصبی مصنوعی شبه‌انسان[۱۱] را برای بازی‌کردن بازی معمای زندانی تکراری معرفی کرده‌اند [۱۵]. مکی[۱۲] کاربرد شبکه‌های عصبی در بازی معمای زندانی را تحت شرایط متفاوتی (مثلاً در نظر گرفتن نویز[۱۳]) در حل مسئله بررسی کرده است [۱۷] و اسگرویی[۱۴] و زیزو[۱۵] نیز کاربرد شبکه‌های عصبی مصنوعی را در بازی‌های ۳×۳ مورد بررسی قرار داده‌اند [۱۸].

در این مقاله شبکه عصبی پادانتشار گراسبرگ[۱۶] برای تحلیل بازی معمای زندانی تکراری که یکی از معروف‌ترین و پرکاربردترین مدل‌های نظریه بازی است، به‌کار گرفته شده است. دلیل استفاده از این شبکه سرعت بالای یادگیری آن است. بدین منظور مقاله در هشت بخش تنظیم شده است: در بخش بعدی به شرح بازی معمای زندانی و کاربردهای آن پرداخته می‌شود. سپس شبکه عصبی پادانتشار گراسبرگ و کاربرد آن در اجرای بازی معمای زندانی تکراری (که رویکرد پیشنهادی این مقاله است) به‌ترتیب در بخش سوم و چهارم تشریح خواهد شد. بخش پنجم نتایج حاصل از شبیه‌سازی را بیان می‌کند. در بخش ششم برای سنجش میزان کارایی شبکه پادانتشار گراسبرگ در اجرای بازی معمای زندانی تکراری، این مسئله با شبکه پرسپترون حل می‌شود و در بخش هفتم نتایج حاصل از اجرای این دو شبکه با نتایج حاصل از به‌کارگیری راهبرد "این به آن در" (TFT)[۱۷] مقایسه می‌شود. بخش هشتم مقاله نیز به نتیجه‌گیری اختصاص یافته است.

۲- بازی معمای زندانی

یکی از معروف‌ترین و مفیدترین بازی‌ها که کاربرد زیادی در اکثر زمینه‌ها به ویژه مسایل اقتصادی و تجاری دارد، بازی معمای زندانی[۱۸] است. این بازی در سال ۱۹۵۰ میلادی توسط درشر[۱۹] و فلوود[۲۰] با هدف توصیف موقعیتی با رفتار معقول (رفتاری که سود قابل انتظار شخص را حداکثر کند) مطرح شد. در این بازی دو متهم به یک اتهام دستگیر شده‌اند. پلیس شواهد کافی برای اثبات جرم متهمان ندارد و به دنبال آن است که دست کم یکی از آنها اعتراف کند. متهمان در سلول‌های جداگانه نگهداری می‌شوند و بنابراین نمی‌توانند با یکدیگر تبادل اطلاعات داشته باشند. در بازجویی زندانیان می‌توانند اعتراف (حالت C) یا سکوت (حالت D) کنند. اگر هیچ کدام اعتراف نکنند، آنها به تخلف کوچکی محکوم و برای یک سال زندانی می‌شوند و اگر تنها یکی از آنها اعتراف کند، آزاد شده و شخص دیگر محکوم به هشت سال زندانی خواهد شد. در نهایت اگر هر دو اعتراف کنند هر دو به پنج سال زندان محکوم می‌شوند [۵].

همچنان که اشاره شد معمای زندانی یکی از مشهورترین و مفیدترین بازی‌ها در اقتصاد است. این مسئله در اقتصاد این گونه بیان می‌گردد: موقعیتی را فرض کنید که دو رقیب در برابر یکدیگر مسابقه‌ای را برگزار می‌کنند که قادر به برقراری ارتباط با هم نیستند. هر یک از دو رقیب دو انتخاب دارند: همکاری (حالت C) یا عدم همکاری (حالت). با توجه به نوع انتخاب رقیبان به هر کدام عایدی تعلق می‌گیرد. اگر هر دو نفر را

1. Imitation
2. Fuzzy Game
3. Chakeri
4. Garazic
5. Cruz
6. Inference
7. Defuzzification
8. Leshno
9. Couraud
10. Liu
11. Like Human Neural Network
12. Macy
13. Noise
14. Sgroi
15. Zizzo
16. Grossberg Counter Propagation
17. Tit for Tat
18. Prison Dilemma
19. Melvin Dresher
20. Merrill Flood

Archive of SID

www.SID.ir

رقیب (بازیکن) اول	رقیب (بازیکن) دوم: همکاری	رقیب (بازیکن) دوم: عدم همکاری
همکاری	۰٫۷،۰٫۷	۰،۱
عدم همکاری	۱،۰	۰٫۳،۰٫۳

شکل ۱: ماتریس بازی معمای زندانی در اقتصاد.

انتخاب کنند، هر کدام ۰/۷ دلار عایدی دریافت می‌کنند. اگر رقیب اول C و رقیب دوم D را انتخاب کند، رقیب اول ۱ دلار عایدی دریافت کرده و دومی هیچ عایدی دریافت نمی‌کند. اگر این حالت برعکس رخ دهد، رقیب دوم ۱ دلار عایدی دریافت کرده و اولی هیچ عایدی دریافت نمی‌کند. در حالتی که هر دو D را انتخاب کنند، به هر کدام ۰/۳ دلار عایدی داده می‌شود. ماتریس این بازی در شکل ۱ نشان داده شده است.

بدیهی است در این بازی راهبرد بهینه برای هر دو بازیکن عدم همکاری است. بازی معمای زندانی اگر به‌صورت دوره‌ای (چندین مرتبه) برگزار شود، به بازی معمای زندانی تکراری[1] تبدیل می‌شود. در این حالت هر بازیکن فرصت می‌کند رفتار خود را با توجه به مشاهداتش از رفتار بازیکن دیگر در مراحل قبل وفق دهد و بدین ترتیب این امکان پدید می‌آید که بازیکن دریابد که آیا طرف مقابل تمایلی به همکاری متقابل دارد یا خیر. نکته حائز اهمیت در این مسئله این است که هر بازیکن به فکر بیشینه‌کردن عایدی خود در طول بازی است و به عایدی بازیکن دیگر اهمیتی نمی‌دهد (هدف بیشینه‌کردن عایدی خود است و نه شکست رقیب). بنابراین انتخاب راهبرد عدم همکاری در طول این بازی تکراری نمی‌تواند راهکار مناسبی برای به‌دست آوردن حداکثر عایدی باشد.

راهبردهایی که در مواجه‌شدن با معمای زندانی تکراری به کار می‌روند به دو دسته تقسیم می‌شوند:

۲-۱ راهبردهای واکنشی[2]

راهبردهایی که بازیکن در هر مرحله، راهبردها و نتایج به‌دست آمده در مراحل قبل را مد نظر قرار داده و طبق آن راهبرد خود را انتخاب می‌کند. از معروف‌ترین و مؤثرترین این راهبردها می‌توان به نمونه‌های زیر اشاره کرد:

الف) این به آن در: در اولین حرکت، راهبرد همکاری را انتخاب و در راهبردهای بعد راهبرد قبل حریف را تقلید می‌کند [۱۹].

ب) اگر بردی بمان، اگر باختی عوض کن[3] (WSLS): اگر عایدی ۰/۷ یا ۱ به‌دست آید، راهبرد قبل تکرار می‌شود و در غیر این صورت راهبرد قبل تغییر می‌یابد [۱۷].

ج) اکثریت: راهبردی که رقیب در بیشتر حالات برگزیده انتخاب شود.

۲-۲ راهبردهای غیر واکنشی[4]

راهبردهایی که در آنها بازیکن بدون توجه به عملکرد رقیب در مراحل قبل راهبرد خود را انتخاب می‌کند. از معروف‌ترین روش‌های این راهبرد می‌توان به نمونه‌های زیر اشاره کرد:

الف) همیشه عدم همکاری $(A\|D)$ یا همیشه همکاری $(A\|C)$.

ب) $CCCD$، DDC، CCD، CD و ... (همکاری و عدم همکاری با الگوی تکراری مشخص).

ج) راهبرد تصادفی.

یک راهبرد واکنشی هوشمندانه، حالت "سوء استفاده" را (یعنی حالتی که بازیکن راهبرد همکاری را انتخاب کند در حالی که حریف عدم همکاری را برگزیند) تحمل نخواهد کرد. بنابراین بهترین عایدی که حریف این راهبرد می‌تواند به آن امیدوار باشد، همکاری دوجانبه است. البته در بازی معمای زندانی تکراری به نظر می‌رسد راهبرد بهینه در راهبردهای واکنشی، راهبرد "این به آن در" است [۱۹] و [۲۰]. در مقابل بهترین راهبرد یک راهبرد غیر واکنشی، "عدم همکاری" است.

این نتایج حاکی از آن است که هیچ یک از راهبردها درباره مسئله بازی معمای زندانی تکراری نمی‌توانند به‌صورت بهینه در مقابل راهبردهای دیگر عمل کند. به همین دلیل استفاده از رویکرد شبکه عصبی برای اتخاذ راهبردهای دیگر در بازی‌کردن معمای زندانی، مناسب به نظر می‌رسد. در این رویکرد سعی بر آن است معماری شبکه‌ای عصبی چنان طراحی شود که به‌جای بهینه‌بودن، کارا و مؤثر بازی کند. در این مقاله شبکه عصبی پادانتشار گراسبرگ برای مدل‌سازی و تحلیل بازی معمای زندانی پیشنهاد شده است، بدین دلیل در بخش بعد به شرح ویژگی‌های این شبکه و چگونگی تحلیل این بازی توسط آن می‌پردازیم.

۳- شبکه عصبی پادانتشار گراسبرگ

شبکه پادانتشار گراسبرگ[5] (GCNN) دو لایه بی‌نظارت[6] و بانظارت[7] معروف به لایه گراسبرگ را با یکدیگر ترکیب می‌کند. معماری این شبکه در شکل ۲ نشان داده شده است. همان‌طور که شکل نشان می‌دهد این شبکه از سه لایه تشکیل شده است: لایه ورودی که الگوهای ورودی را می‌خواند و آنها را به سوی شبکه می‌فرستد، لایه نهانی که همان لایه بی‌نظارت است، به‌صورت رقابتی[8] کار می‌کند و هر یک از الگوهای ورودی را به تمامی نرون‌های نهانی متصل می‌کند و لایه خروجی به لایه گراسبرگ مشهور است [۲۱] و [۲۲].

الگوریتم یادگیری شبکه پادانتشار گراسبرگ را می‌توان در دو مرحله شرح داد: در مرحله اول وزن پیوندهای بین لایه ورودی و لایه بی‌نظارت به‌روز می‌شود، در حالی که در مرحله دوم وزن بین لایه بی‌نظارت (رقابتی) و لایه گراسبرگ روزآمد می‌شوند. با جزئیات بیشتر می‌توان این الگوریتم را به‌صورت زیر شرح داد:

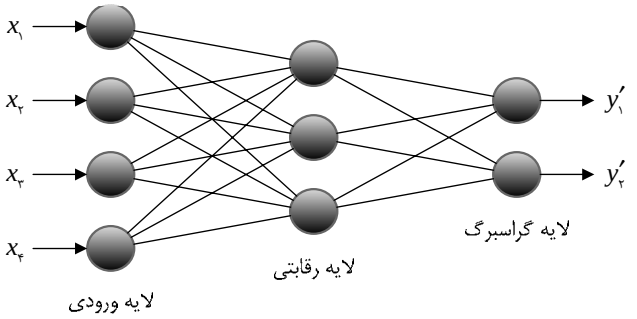
فرض کنید بردار x به شبکه اعمال شود و t خروجی مطلوب مربوط به این بردار باشد. بردار y، بردار خروجی واقعی شبکه مربوط به بردار ورودی x است. هدف کمینه‌کردن اختلاف بین t و y است. فرض کنید $W_j = [w_{ji}]$ بردار وزن‌های اولیه لایه رقابتی برای نرون j ام در این لایه باشد که w_{ji} وزن پیوند بین نرون j ام در لایه رقابتی و نرون i ام در لایه ورودی باشد. در این صورت در لایه رقابتی مقدار زیر برای تمامی نرون‌ها محاسبه می‌شود

$$S_j = \sum_{i=1}^{m} w_{ij} x_i \qquad (1)$$

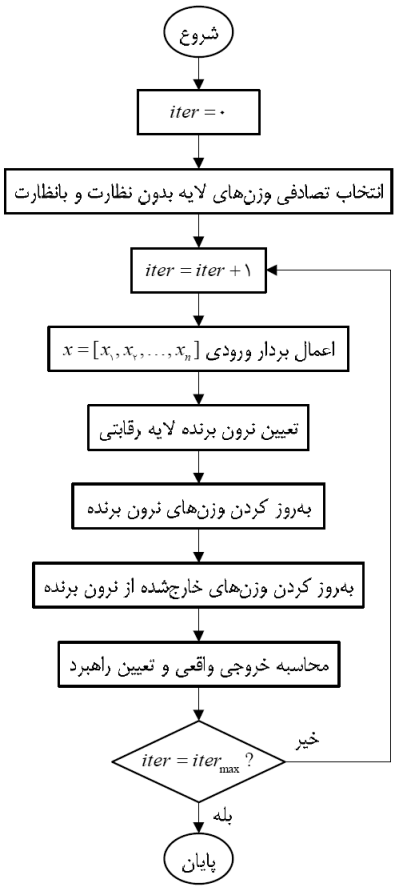
1. Iterated Prison Dilemma
2. Responsive Strategies
3. Win - Stay, Lose - Shift
4. Nonresponsive Strategies
5. Grossberg Counter - Propagation Neural Network
6. Unsupervised
7. Supervised
8. Competitive

Archive of SID

www.SID.ir



شکل ۲: معماری شبکه پادانتشار گراسبرگ.



شکل ۳: الگوریتم یادگیری شبکه پادانتشار گراسبرگ.

در این رابطه i معرف نرون iام لایه ورودی، j معرف نرون j ام لایه رقابتی و m تعداد نرون‌های لایه ورودی است و نرونی که بیشترین مقدار S_j را دارد، به‌عنوان نرون برنده معرفی می‌شود. در نتیجه این رقابت، خروجی نرون برنده برابر ۱ و خروجی دیگر نرون‌ها برابر ۰ در نظر گرفته می‌شود. بنابراین خروجی نرون j ام در لایه رقابتی را می‌توان به‌صورت زیر نشان داد

$$z_j = \begin{cases} 1 & \text{if } j \text{ is the winner} \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

پس از این مرحله، وزن مربوط به نرون j ام در لایه رقابتی و نرون iام در لایه ورودی به‌صورت زیر روزآمد می‌شود

$$w_{ji}^n = w_{ji}^o + \alpha \times (x_i - w_{ji}^o) \quad , \quad 0 < \alpha \ll 1 \qquad (3)$$

که α آهنگ یادگیری است.

در لایه بانظارت (لایه گراسبرگ) وزن‌های خارج‌شده از نرون برنده بر اساس قاعده ویدرو- هاف[1] (Mira ،۲۰۰۷) اصلاح می‌شوند

$$v_j^n = v_j^o + \beta \times (t - y) \quad , \quad 0 < \beta < 1 \qquad (4)$$

که β معروف به آهنگ یادگیری گراسبرگ[2] است.

با در نظر گرفتن تابع سیگموید[3] در خروجی شبکه داریم

$$y = \frac{1}{1 + e^{-v}} \qquad (5)$$

در کل می‌توان گفت فرایند یادگیری، بردار ورودی را با بردار خروجی بر اساس دو الگوریتم مشهور به یکدیگر متصل می‌کند: نگاشت خودسازمان‌ده کوهنن[4] برای پیداکردن شبیه‌ترین بردار آموزشی و نگاشت گراسبرگ برای پیش‌بینی بردار خروجی [۲۳]. الگوریتم یادگیری شبکه پادانتشار گراسبرگ در طول بازی معمای زندانی تکراری که تعداد چرخه (دست یا دور) از پیش تعیین شده است، در شکل ۳ نشان داده شده است. در این شکل $iter$ یک شمارنده است و بازی تا جایی ادامه می‌یابد که شمارنده به $iter_{\max}$ (تعداد دست از پیش تعیین شده) برسد.

٤- مدل‌سازی و تحلیل بازی معمای زندانی به کمک شبکه پادانتشار گراسبرگ

برای طراحی شبکه‌ای که قادر به انجام بازی باشد، باید بار دیگر اجزای بازی از جمله راهبرد و عایدی را بررسی و الگوهای ورودی و خروجی‌های متناظر با آنها را مشخص کنیم. تصمیم‌گیرنده قبل از تصمیم‌گیری راهبردهای پیش روی خود را می‌داند و با توجه به بررسی راهبردهای مراحل قبل حریف و عایدی‌های آن راهبرد بهینه را انتخاب می‌کند، گویی مسئله‌ای از جنس بهینه‌سازی را حل می‌کند. برای تحلیل بازی، راهبردها و عایدی‌های بازیکنان در مراحل قبل را به‌عنوان ورودی شبکه در نظر می‌گیریم. برای خروجی شبکه نیز می‌توان قراردادی در نظر گرفت. برای مثال اگر دو راهبرد وجود دارد و خروجی شبکه بین ۰ و ۱ باشد، مقادیر بین ۰ و ۰/۵ نشان‌دهنده یکی از راهبردها و مقادیر بین ۰/۵ و ۱ نشان‌دهنده راهبرد دیگر است.

در این مسئله ورودی شبکه عصبی، نتیجه دوره‌های قبل است که نوع راهبرد و عایدی هر بازیکن را شامل می‌شود. بدین ترتیب اگر نتایج یک دوره قبل به‌عنوان ورودی در نظر گرفته شوند، ۴ ورودی (راهبرد و عایدی بازیکن اول، راهبرد و عایدی بازیکن دوم) و اگر نتایج دو دوره قبل نیز در نظر گرفته شوند، در مجموع ۸ ورودی خواهیم داشت. به همین ترتیب با افزایش تعداد دورهای قبل، تعداد ورودی‌های شبکه افزایش می‌یابد. نکته حائز اهمیت این است که مقدار واقعی تولیدشده توسط شبکه در مرحله قبل به‌عنوان ورودی کنونی در نظر گرفته می‌شود.

تابع فعال‌سازی لایه گراسبرگ، تابع سیگموید $1/(1+e^{-x})$ است، بنابراین خروجی شبکه عصبی مقداری در بازه $[0,1]$ است که نوع راهبرد (همکاری یا عدم همکاری) در دوره مربوطه را نشان می‌دهد. این مقدار اگر بزرگ‌تر از ۰/۵ باشد راهبرد همکاری (C) و در غیر این صورت راهبرد عدم همکاری (D) است. در این شبکه هر واحد (نرون) به همه واحدهای لایه (های) مجاور متصل است و وزن‌های اولیه به‌طور یکنواخت در بازه $[0,1]$ انتخاب شده است.

1. Widrow - Hoff Rule
2. Grossberg Learning Rate
3. Sigmoidal Function
4. Self - Organizing Kohonen

Archive of

www.SID.ir

جدول ۱: میانگین عایدی حاصل از شبکه پادانتشار گراسبرگ با معماری‌های متفاوت در مقابل سایر راهبردها (۱۰۰۰ تکرار).

راهبرد / شبکه	TFT	WSLS	$A\|\|D$	$A\|\|C$	CD	CCD	DDC	$CCCD$	$DDDC$	$DDDDC$	تصادفی
GCNN (۴,۳,۱)	۰/۷۰	۰/۷۰	۰/۲۷	۰/۷۰	۰/۵۲	۰/۶۰	۰/۴۷	۰/۶۷	۰/۴۲	۰/۳۹	۰/۵۳
GCNN (۴,۵,۱)	۰/۷۰	۰/۷۰	۰/۲۷	۰/۷۰	۰/۵۰	۰/۶۰	۰/۴۶	۰/۶۷	۰/۴۲	۰/۳۸	۰/۵۲
GCNN (۴,۸,۱)	۰/۷۰	۰/۶۴	۰/۲۷	۰/۷۰	۰/۵۱	۰/۶۰	۰/۴۷	۰/۶۷	۰/۴۲	۰/۳۹	۰/۵۱
GCNN (۴,۱۰,۱)	۰/۷۰	۰/۶۹	۰/۲۷	۰/۷۰	۰/۵۰	۰/۵۸	۰/۴۶	۰/۶۷	۰/۴۲	۰/۳۹	۰/۵۲
GCNN (۸,۳,۱)	۰/۷۰	۰/۷۰	۰/۲۷	۰/۷۰	۰/۵۲	۰/۶۲	۰/۴۷	۰/۶۷	۰/۴۲	۰/۳۹	۰/۵۳
GCNN (۸,۵,۱)	۰/۷۰	۰/۷۰	۰/۲۷	۰/۷۰	۰/۴۸	۰/۶۱	۰/۴۴	۰/۶۷	۰/۴۲	۰/۳۹	۰/۵۲
GCNN (۸,۸,۱)	۰/۷۰	۰/۷۰	۰/۲۷	۰/۷۰	۰/۴۹	۰/۶۳	۰/۴۵	۰/۶۷	۰/۴۲	۰/۳۹	۰/۵۱
GCNN (۸,۱۰,۱)	۰/۷۰	۰/۷۰	۰/۲۷	۰/۷۰	۰/۴۹	۰/۶۳	۰/۴۴	۰/۶۷	۰/۴۰	۰/۳۸	۰/۵۲

بدین دلیل که راهبرد بازیکن مقابل برای شبکه عصبی مشخص نیست و حریف می‌تواند هر راهبرد دلخواهی را مد نظر قرار دهد، در نتیجه یادگیری شبکه بی‌درنگ[1] صورت می‌پذیرد. ضمناً از آنجا که بازی معمای زندانی تکراری در چرخه‌های زیادی انجام می‌شود، فرض بر این است که شبکه در آغاز بازی به‌طور تصادفی راهبرد خود را انتخاب می‌کند و در مراحل آغازین با تنظیم وزن‌های خود سعی می‌کند بیشترین عایدی ممکن را از حریف بگیرد. به‌منظور سرعت‌دادن به فرایند یادگیری مقدار آهنگ یادگیری[2] (β) نسبتاً بزرگ انتخاب می‌شود تا شبکه به سرعت وفق داده شود و بعد تدریجاً کاهش می‌یابد. در لایه رقابتی، وزن‌های نرون برنده و در لایه گراسبرگ، وزن‌های خارج‌شده از نرون برنده (بر اساس قاعده ویدرو- هاف) اصلاح می‌شوند.

بدین دلیل که شبکه از قبل نمی‌داند کدام راهبرد به عایدی بهتری می‌انجامد، لذا باید تعیین مقدار هدف به‌صورت شهودی[3] صورت پذیرد. بدین منظور از دو روش می‌توان بهره جست: روش شهود همکارانه[4] و روش شهود قدردانانه[5] [۲۴]. در روش همکارانه اگر رقیب بیشتر همکاری کرده باشد (یعنی در مرحله قبل C را انتخاب کرده باشد)، راهبرد قبل تکرار می‌شود و در غیر این صورت راهبرد کنونی عکس راهبرد قبل انتخاب می‌گردد. زیرا از آنجا که یک بازیکن همیشه بهترین راهبرد را برمی‌گزیند زمانی که دیگری همکاری کند، شبکه باید همکاری‌کردن را به حریف خود تحمیل (القا) کند. این روش در اصل نوعی یادگیری با نظارت است که در آن مقدار مطلوب مستقیماً به وسیله نتیجه مشخص می‌گردد. در مقابل در روش قدردانانه اصل بر این است که هر چه عایدی بیشتر شود، همکاری هم بیشتر می‌شود.

با اندکی تأمل در می‌یابیم عملکرد روش اول در برابر راهبردهای غیر واکنشی خوش‌بینانه است، زیرا این روش فرض می‌کند اقدام فرد بر تصمیم حریف تأثیر خواهد داشت، در صورتی که همیشه این گونه نیست. از سوی دیگر در روش دوم اگر راهبرد نخست بازیکن عدم همکاری باشد، تمایلی به همکاری دوجانبه (هر دو همکاری کنند) در برابر راهبردهای واکنشی وجود ندارد. به همین دلیل مناسب‌ترین رویکرد، استفاده از ترکیب خطی دو روش فوق به‌صورت زیر است

$$t = t_c \times f + t_g \times (1 - f) \qquad (۶)$$

که در آن $f \in [0,1]$ عددی ثابت و t_c، t_g و t به‌ترتیب مقادیر مطلوب روش همکارانه، روش قدردانانه و ترکیب این دو روش (مقدار مطلوب اصلی) هستند [۲۴].

۵- نتایج شبیه‌سازی

همچنان که در بخش ۲ توضیح داده شد، بازی معمای زندانی تکراری با دو بازیکن اجرا می‌شود که برای تحلیل آن به کمک شبکه پادانتشار گراسبرگ، یکی از بازیکنان این شبکه با معماری توضیح داده شده در بخش ۳ به‌عنوان بازیکن اول و در طرف دیگر بازیکن دوم قرار دارد. بازیکن دوم با راهبردهای گوناگون اعم از واکنشی و غیر واکنشی مقابل شبکه بازی می‌کند. شایان ذکر است که هر بازی از ۱۰۰۰ چرخه (دست) تشکیل شده (با این تعداد چرخه که نه بیش از حد طولانی و نه کوتاه است می‌توان هم عملکرد شبکه در مدت زمانی متوسط را بررسی کرد و هم در زمانی که برای یادگیری شبکه نیاز است به عایدی معقولی دست یافت) و فرض بر این است که در همه بازی‌ها، شبکه کار خود را با عدم همکاری (D) آغاز می‌کند. این فرض بدان دلیل است که در غیر این صورت (یعنی با فرض این که شبکه عصبی با همکاری بازی را شروع می‌کند)، شبکه فوراً راهبرد بهینه همکاری دوطرفه (هر دو همکاری کنند) را بر می‌گزینند.

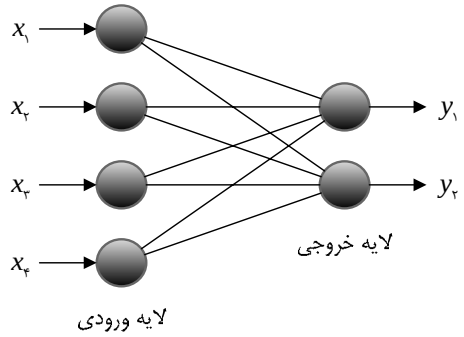
شبکه طراحی‌شده دارای یک نرون خروجی است که اگر خروجی آن مقداری بزرگ‌تر از ۰/۵ باشد "راهبرد همکاری" و در غیر این صورت "راهبرد عدم همکاری" به‌عنوان راهبرد شبکه انتخاب می‌شود. شبکه مورد نظر با معماری‌های متفاوت مورد بررسی قرار گرفت و بهترین معماری آن در نظر گرفته می‌شود. با توجه به نوع مسئله تابع فعالیت تمامی شبکه‌های طراحی‌شده، تابع سیگموید است و تفاوت این شبکه‌ها تنها در تعداد نرون‌های ورودی و نهان است. اگر شبکه با چهار ورودی ساخته شود، راهبرد بازیکن اول در مرحله قبل، عایدی بازیکن اول در مرحله قبل، راهبرد بازیکن دوم در مرحله قبل و عایدی آن بازیکن در مرحله قبل به‌عنوان ورودی‌های شبکه در نظر گرفته می‌شوند. با در نظر گرفتن هشت ورودی نتایج دو مرحله قبل نیز به‌عنوان ورودی در نظر گرفته می‌شوند. با توجه به نکات فوق عملکرد شبکه با معماری‌های متفاوت در بازی معمای زندانی تکراری با ۱۰۰۰ تکرار در مقابل راهبردهای متفاوت معرفی‌شده در جدول ۱ آمده است. ستون اول این جدول معماری‌های متفاوت شبکه را نمایش می‌دهد. برای مثال GCNN(۴,۱۰,۱) نمایانگر شبکه‌ای با چهار نرون ورودی، ده نرون نهانی و یک نرون خروجی است.

نتایج نشان می‌دهد شبکه گراسبرگ با هشت ورودی و سه نرون در لایه رقابتی نسبت به سایر معماری‌های تحقیق‌شده عملکرد بهتری دارد.

1. On Time
2. Learning Rate
3. Heuristic
4. Cooperativeness Heuristic
5. Gratitude Heuristic

Archive of SID

www.SID.ir



شکل ۴: ساختار شبکه پرسپترون.

البته برای مثال هنگامی که شبکه در مقابل راهبرد *CCCD* قرار دارد وقتی تعداد چرخه‌های بازی را به ۵۰۰ کاهش می‌دهیم، خطای روش از ۰/۰۰۰۴ به ۰/۲ و میانگین عایدی از ۰/۶۷ به ۰/۵۴ تغییر می‌یابد و با افزایش تعداد چرخه‌ها، میانگین عایدی و خطا تغییر محسوسی نکرده و تنها سرعت همگرایی پایین می‌آید و به همین علت تعداد چرخه‌ها را همان ۱۰۰۰ در نظر می‌گیریم.

در این مقاله برای سنجش میزان کارایی شبکه پادانتشار گراسبرگ در انجام بازی معمای زندانی تکراری، این بازی را به کمک شبکه عصبی پرسپترون نیز اجرا می‌کنیم. از این رو در بخش بعد به شرح شبکه پرسپترون و تحلیل بازی به کمک آن می‌پردازیم و نتایج حاصل از اجرای بازی توسط این دو شبکه را با حالتی که یکی از بازیکنان در مقابل هر دو شبکه و راهبردهای پیشنهادشده در بازی معمای زندانی تکراری، راهبرد "این به آن در" (TFT) را اتخاذ می‌کند، مقایسه می‌کنیم.

۶- مقایسه توانمندی شبکه گراسبرگ در مقایسه با شبکه پرسپترون در اجرای بازی

شبکه پرسپترون، شبکه‌ای پیش‌رو، تک‌لایه و دودویی است، یعنی خروجی شبکه برابر ۱ یا ۰ است و از یک لایه ورودی و یک لایه خروجی تشکیل شده است. معماری این شبکه در شکل ۴ نشان داده شده است. تابع فعالیت[1] هر نرون، تابع دودویی نامتقارن است. در این شبکه فقط وزن‌های متصل به نرون‌های خروجی قابل تنظیم هستند. ممکن است یک نرون ثابت به غیر از نرون‌های دیگر داشته باشیم که به آن نرون بایاس می‌گویند و مقدارش ۰ یا ۱ است.

الگوریتم آموزش شبکه پرسپترون به‌صورت زیر است:

۱) بردار وزن پیوندهای شبکه W در مرحله اول به‌طور تصادفی انتخاب می‌شود. فرض کنید $W_j=[w_{ji}]$ بردار وزن پیوندهای بین نرون j ام و تمامی نرون‌های لایه خروجی باشد.

۲) الگوهای ورودی نمونه $X=[x_1,x_2,\ldots,x_n]$ به شبکه اعمال می‌شوند. به ازای هر الگوی ورودی یک الگوی خروجی مطلوب $Y=[y_1,y_2,\ldots,y_m]$ وجود دارد.

۳) خروجی واقعی نرون j ام در مرحله t از رابطه زیر محاسبه می‌شود

$$y'_j(t)=f[\sum_{i=0}^{n} w_{ij}(t).x_i(t)] \quad , \quad j=1,2,\ldots,n \qquad (۷)$$

که y_j خروجی واقعی شبکه، x_i ورودی شبکه، w_{ij} وزن بین نرون i ام و نرون j ام، j تعداد خروجی و f تابع فعالیت است.

۴) وزن پیوندهای شبکه طبق قاعده ویدرو- هاف روزآمد می‌شود. بدین منظور می‌توان نوشت (به (۸) پرانتز اضافه شد)

$$w_{ij}(t+1)=w_{ij}(t)+\alpha.(y(t)-y'(t)).x_i \quad , \quad 0<\alpha<1 \qquad (۸)$$

که α آهنگ یادگیری است.

۵) اضافه‌کردن یک واحد به t و تکرار الگوریتم از مرحله ۲ تا جایی که بردار وزن شبکه به تثبیت برسد و تغییرات ناچیز داشته باشد [۲۵].

در رابطه با این مسئله تابع فعالیت شبکه پرسپترون نیز تابع سیگموید $1/(1+e^{-x})$ است. آموزش در این شبکه نیز به‌صورت برخط صورت می‌پذیرد و مقدار آهنگ یادگیری α بزرگ (در حدود ۱) در نظر گرفته می‌شود تا شبکه به سرعت وفق داده شود. الگوهای ورودی و خروجی مطلوب اعمال‌شده به این شبکه شبیه به الگوهای ورودی به شبکه پادانتشار گراسبرگ است که در بخش پنجم توضیح داده شد. مزیت عمده شبکه پادانتشار گراسبرگ بر پرسپترون، سرعت آن است و علت بالاتربودن سرعت شبکه پادانتشار گراسبرگ همان طور که در بخش ۳ گفته شد این است که این شبکه به علت وجود لایه نهانی که همان لایه بی‌نظارت (رقابتی) است و در همان ابتدا نورون‌های برنده را انتخاب می‌کند، بدون نظارت یاد می‌گیرد. اما در شبکه پرسپترون از آنجا که این شبکه درصدد کمینه‌کردن خطا بین اندازه خروجی شبکه و خروجی مطلوب است، یادگیری بانظارت است که این موضوع باعث لختی و سرعت پایین پرسپترون می‌شود [۲۶]. تفاوت بارز این دو شبکه در این شبکه لایه رقابتی است و سرعت شبکه پادانتشار گراسبرگ را به نسبت پرسپترون که فاقد این لایه است بالا می‌برد.

معماری شبکه پرسپترون از لحاظ ورودی و خروجی شبیه شبکه گراسبرگ است. عملکرد شبکه عصبی پرسپترون با دو معماری متفاوت در مقابل راهبردهای متفاوت در جدول ۲ آورده شده است. همچنان که ملاحظه می‌شود در تحلیل بازی به کمک پرسپترون در دو مورد یعنی حالت‌های *CCD* و *CCCD* نتایج پرسپترون با هشت نرون ورودی (PNN(۸,۱)) بهتر از پرسپترون با چهار نرون ورودی (PNN(۴,۱)) است اما در بقیه موارد (PNN(۴,۱)) نتایج بهتری را به‌دست می‌دهد. در نتیجه بهتر است تنها از نتایج به‌دست آمده از یک دوره قبل به‌عنوان ورودی استفاده کنیم. به‌عبارت دیگر پرسپترون با چهار نرون ورودی و یک نرون خروجی نتایج بهتری نسبت به پرسپترون با معماری هشت نرون ورودی و یک نرون خروجی می‌دهد.

برای میزان سنجش کارایی شبکه گراسبرگ و پرسپترون، عملکرد راهبرد "این به آن در" نیز در جدول ۳ نشان داده شده است. شایان ذکر این که در بسیاری از تحقیقات، این راهبرد به‌عنوان مؤثرترین و بهترین راهکار در اجرای بازی معمای زندانی تکراری معرفی شده است [۱۹ و ۲۰].

در جدول ۴ مقایسه عملکرد بین شبکه پرسپترون، گراسبرگ و راهبرد TFT نشان داده شده است. نتایج به‌دست آمده نشان از عملکرد مطلوب شبکه گراسبرگ دارد. شبکه گراسبرگ در برابر راهبردهای واکنشی عملکرد برابری با TFT دارد، در حالی که شبکه پرسپترون در برابر راهبرد WSLS عملکرد ضعیف‌تری نسبت به گراسبرگ و TFT داراست. اگرچه شبکه گراسبرگ در برابر راهبرد $A\|D$ عایدی کمتری نسبت به TFT را به‌دست می‌دهد، اما در برابر سایر روش‌ها میزان عایدی دریافتی شبکه گراسبرگ بیشتر یا مساوی TFT است. میزان عایدی دریافتی شبکه گراسبرگ نسبت به TFT در راهبردهای *CCD* و *CCCD* افزایش قابل ملاحظه‌ای دارد. از سوی دیگر شبکه عصبی پرسپترون در الگوهای

1. Activation Function

Archive of SID

www.SID.ir

جدول ۲: میانگین عایدی حاصل از شبکه پرسپترون با دو معماری متفاوت در مقابل سایر راهبردها (۱۰۰۰ دست).

راهبرد	TFT	WSLS	$A\|\|D$	$A\|\|C$	*CD*	*CCD*	*DDC*	*CCCD*	*DDDC*	*DDDDC*	تصادفی
PNN (٤,١)	۰٫۷۰	۰٫۶۹	۰٫۲۹	۰٫۷۰	۰٫۴۸	۰٫۵۲	۰٫۴۳	۰٫۵۴	۰٫۴۳	۰٫۴۰	۰٫۵۲
PNN (٨,١)	۰٫۷۰	۰٫۶۹	۰٫۲۸	۰٫۷۰	۰٫۴۸	۰٫۵۴	۰٫۴۰	۰٫۵۷	۰٫۳۹	۰٫۳۸	۰٫۵۰

جدول ۳: میانگین عایدی حاصل از TFT در مقابل سایر راهبردها (۱۰۰۰ دست).

راهبرد	ANN	WSLS	$A\|\|D$	$A\|\|C$	*CD*	*CCD*	*DDC*	*CCCD*	*DDDC*	*DDDDC*	تصادفی
TFT	۰٫۷۰	۰٫۷۰	۰٫۳۰	۰٫۷۰	۰٫۵۰	۰٫۵۷	۰٫۴۳	۰٫۶۰	۰٫۴۰	۰٫۳۸	۰٫۵۰

جدول ۴: مقایسه میانگین عایدی حاصل از شبکه عصبی پادانتشار گراسبرگ، پرسپترون و راهبرد TFT در مقابل سایر راهبردها (۱۰۰۰ دست).

راهبرد	TFT	WSLS	$A\|\|D$	$A\|\|C$	*CD*	*CCD*	*DDC*	*CCCD*	*DDDC*	*DDDDC*	تصادفی
GCNN (٨,٣,١)	۰٫۷۰	۰٫۷۰	۰٫۲۷	۰٫۷۰	**۰٫۵۲**	**۰٫۶۲**	**۰٫۴۷**	**۰٫۶۷**	۰٫۴۲	۰٫۳۹	**۰٫۵۳**
PNN (۴,١)	۰٫۷۰	۰٫۶۹	۰٫۲۹	۰٫۷۰	۰٫۴۸	۰٫۵۲	۰٫۴۳	۰٫۵۴	**۰٫۴۳**	**۰٫۴۰**	۰٫۵۲
TFT	۰٫۷۰	۰٫۷۰	۰٫۳۰	۰٫۷۰	۰٫۵۰	۰٫۵۷	۰٫۴۳	۰٫۶۰	۰٫۴۰	۰٫۳۸	۰٫۵۰

DDC، *DDDC* و تصادفی بر TFT غلبه می‌کند. این شبکه در برابر الگوهای *DDC*، *DDDC* و $A\|\|D$ بر شبکه گراسبرگ برتری دارد. بدین ترتیب ملاحظه می‌شود با توجه به نتایج به‌دست آمده بهترین عملکرد از آنِ شبکه گراسبرگ است. دلیل عملکرد مطلوب این شبکه را می‌توان لایه رقابتی، سرعت بالای یادگیری و پاسخ آن دانست.

۷- نتیجه‌گیری

در این مقاله عملکرد شبکه عصبی در حل بازی معمای زندانی تکراری مورد مطالعه قرار گرفت. بدین منظور شبکه عصبی گراسبرگ با قابلیت یادگیری سریع طراحی شد. به‌منظور بررسی عملکرد شبکه عصبی گراسبرگ نتایج حاصل از شبیه‌سازی با نتایج حاصل از شبکه عصبی پرسپترون و راهبرد TFT مقایسه گردید. مقایسه نتایج به‌دست آمده نشان از برتری شبکه گراسبرگ نسبت به شبکه پرسپترون و TFT دارد، به گونه‌ای که این شبکه قادر است به نحو مؤثرتری نسبت به شبکه عصبی پرسپترون و راهبرد TFT بازی معمای زندانی تکراری را انجام داده و عایدی بیشتری دریافت کند. از این رو شبکه عصبی گراسبرگ می‌تواند به‌عنوان ابزاری مناسب برای مدل‌سازی بازی معمای زندانی تکراری به کار گرفته شود. از سوی دیگر با توجه به آن که مسایل بسیاری از جمله رقابت بازار در بورس [۲۷]، قیمت‌گذاری [۵]، کنترل حرکت روبات‌ها [۲۸]، امنیت در شبکه‌های اقتضایی [۶] و ... به کمک بازی معمای زندانی تحلیل می‌شود، از این رو شبکه عصبی گراسبرگ می‌تواند روش مناسبی برای مدل‌سازی این مسایل در قالب بازی باشد.

مراجع

[1] S. C. Sharma, *Operation Research*, New Dehli: Discovery Publishing House, 2006.

[2] I. K. Geckil and P. Anderson, *Applied Game Theory and Strategic Behavior*, London: Chapman & Hall/CRC, 2010.

[3] R. Ginevicius and A. Krivka, "Application of game theory for duopoly market analysis," *J. of Business Economics and Management*, vol. 9, no. 3, pp. 207-217, Oct. 2007.

[4] K. W. Lye and J. Wing, "Game strategies in network security," *Int. J. of Information Security*, vol. 4, no. 1-2, pp. 71-86, Feb. 2005.

[5] J. Leino, *Applications of Game Theory in Ad Hoc Networks*, in Department of Engineering Physics and Mathematics, Helsinki University of Technology, p. 67, 2003.

[6] A. B. Mackenzie and L. A. Dasilva, *Game Theory for Wireless Engineers*, Morgan & Claypool Publishers, 2006.

[7] M. Leng and M. Parlar, Game Theory Applications in Supply Chain Management: a Review, Infor, p. 220, 2005.

[8] S. Moretti, "Game theory applied to gene expression analysis," *4OR: A Quarterly Journal of Operations Research*, vol. 7, no. 2, pp. 195-198, 2009.

[9] Y. Shoham and K. Leyton - Brown, Multiagent Systems Algorithmic, Game-Theoretic, and Logical Foundations, Cambridge University Press, 2009.

[10] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*, London, England: The MIT Press Cambridge, Massachusetts, 1994.

[11] T. Rees, An Introduction to Evolutionary Game Theory, pp. 1-4, 2005.

[12] N. Nisan, T. Roughgarden, E. Tardos, and V. Vizinari, *Algorithmic Game Theory*, Cambridge University Press, 2007.

[13] A. Chakeri, A. N. Dariani, and C. Lucas, "How can fuzzy logic determine game equilibriums better?" in *Proc. Int. IEEE Conf. Intelligent Systems*, vol. 2, pp. 51-56, 6-8 Sep. 2008.

[14] D. Garagic and J. B. Cruz, "An approach to fuzzy noncooperative nash games," *Optimization Theory and Application*, vol. 118, no. 3, pp. 475-491, Sep. 2003.

[15] B. Couraud and P. Liu, "Use of neural networks as decision makers in strategic situations," in *Proc. of the 8th Int. Conf. on Machine Learning and Cybernetics*, vol. 3, pp. 1280-1285, 12-15 Jul. 2009.

[16] M. Leshno, D. Moller, and P. Ein - Dor, "Neural nets in a group decision process," *Int J. Game Theory*, vol. 31, no. 3, pp. 447-467, Jun. 2003.

[17] M. Macy, "Natural selection and social learning in prisoner's dilemma: coadaptation with genetic algorithms and artificial neural networks," *Sociological Methods and Research*, vol. 25, no. 1, pp. 103-137, Aug. 1996.

[18] D. Sgroi and D. J. Zizzo, "Learning to play 3×3 games: neural networks as bounded-rational players," *J. of Economic Behavior and Organization*, vol. 69, no. 1, pp. 27-38, Jan. 2009.

[19] R. Axelrod, *The Evolution of Cooperation*, New York: Basic Books, 1984.

[20] R. Axelrod, "More effective choice in the iterated prisoner's dilemma," *Conflict Resolution*, vol. 24, no. 3, pp. 379-403, Sep. 1980.

[21] R. Hetcht-Nielsen, "Counterpropagation networks," *Applied Optics*, vol. 26, no. 23, pp. 4979-4984, 1987.

[22] S. C. Juang, Y. S. Tarng, and H. R. Lii, "A comparison between the back-propagation and counter-propagation networks in the modeling of the TIG welding process," *J. of Materials Processing Technology*, vol. 75, pp. 54-62, 1998.

[23] A. I. Margaris and E. Kotsialos, "Parallel counter-propagation networks," in *Proc. of the Int. Conf. on Theory and Applications of Mathematics and Informatics, ICTAMI*, pp. 306-324, Thessaloniki, Greece, 2004.

[24] J. O'Madadhain, Neural Network-Based Strategies for the Iterated Prisoner's Dilemma, Winter 2002.

[25] S. Haykin, Neural Networks a Comprehensive Foundation, Prentice Hall, 1999.

[26] L. W. Freriks, P. J. M. Cluitmans, and M. J. van Gils, *The Adaptive Resonance Theory Network: (Clustering-) Behaviour in Relation with Brainstem Auditory Evoked Potential Patterns*, Eindhoven University of Technology Research Reports, Netherlands, p. 113, 1992.

Archive of SID

www.SID.ir

[27] J. D. Miller, Game Theory at Work: How to Use Game Theory to Outthink and Outmaneuver Your Competition, New York: McGraw-Hill, 2003.
[28] A. Birk and J. Wiernik, "An N-player prisoner's dilemma in a robotic ecosystem," *Int. J. Robotics and Autonomous Systems*, vol. 39, pp. 223-233, 2002.

غلامعلی منتظر در سال ۱۳۴۸ در کازرون(فارس) به دنیا آمد. او در سال ۱۳۷۰ مدرک کارشناسی خود را در رشتهٔ مهندسی برق از دانشگاه صنعتی خواجه نصیرالدین طوسی و سپس در سال‌های ۱۳۷۳ و ۱۳۷۷ مدارک کارشناسی ارشد و دکتری خود را در همین رشته از دانشگاه تربیت مدرّس اخذ کرد.

پس از اتمام تحصیل به عضویت هیأت علمی دانشگاه تربیت مدرّس درآمد و در حال حاضر دانشیار مهندسی فنّاوری اطلاعات در این دانشگاه است. حوزه‌های تخصصی وی شامل نرم‌رایانش (نظریهٔ مجموعه‌های فازی، شبکه‌های عصبی مصنوعی، نظریهٔ مجموعه‌های نادقیق) و کاربرد آن در سیستم‌های اطلاعاتی (همچون سیستم یادگیری الکترونیکی و تجارت الکترونیکی) است. وی تاکنون بیش از ۷۰ مقاله در نشریات معتبر علمی و بیش از ۱۵۰ مقاله در کنفرانس‌های معتبر علمی ملّی و بین‌المللی منتشر کرده است. علاوه بر این حائز دریافت جوایز معتبر علمی از جمله برگزیدهٔ جشنوارهٔ بین‌المللی خوارزمی، برندهٔ کتاب سال دانشگاهی ایران، پژوهشگر برگزیدهٔ آیسسکو و متخصص برجسته فنّاوری اطلاعات ایران شده است.

نجمه رستگار رامشه در سال ۱۳۸۵ تحصیلات خود را در مقطع کارشناسی ریاضی کاربردی دردانشگاه گیلان به پایان رساند و در حال حاضر دانشجوی کارشناسی ارشد رشتهٔ مهندسی فناوری اطلاعات دانشگاه تربیت مدرس است. زمینه‌های علمی مورد علاقه نام‌برده عبارتند از: نظریهٔ مجموعه‌های فازی، شبکه‌های عصبی مصنوعی و کاربرد آن در سیستم‌های اطلاعاتی همچون سیستم یادگیری الکترونیکی.

علیرضا عسکرزاده تحصیلات خود را در مقطع کارشناسی در سال ۱۳۸۶ از دانشگاه شهید باهنر کرمان و مقاطع کارشناسی ارشد و دکتری را به ترتیب در سال‌های ۱۳۸۸ و ۱۳۹۱ از دانشگاه شهید بهشتی در رشته مهندسی برق گرایش قدرت به پایان رسانده و هم‌اکنون عضو هیأت علمی دانشگاه تحصیلات تکمیلی صنعتی و فنّاوری پیشرفته کرمان است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: هوش مصنوعی، بهینه‌سازی سیستم‌های قدرت و انرژی‌های تجدیدپذیر.

Archive of SID

www.SID.ir